Pay-for-Performance Initiatives: Modest Benefits for Improving Healthcare Quality

Amit Sura, MD, MBA; Nirav R. Shah, MD, MPH

Amit Sura

Background: Pay-for-performance initiatives have been suggested as a way to improve the quality of patient care and provide incentives to improve providers' performance. The Centers for Medicare & Medicaid Services has endorsed such programs to improve quality of care.

Objective: To examine the state of quality initiatives endorsed by the Centers for Medicare & Medicaid Services in which institutions, provider groups, and physicians are awarded incentives based on adherence to composite metrics.

Method: A literature search was conducted using the keywords "pay-for-performance," "quality improvement," "medical errors," and "physician incentive plans."

Results: Although quality of care has improved in healthcare settings that engage in pay-for-performance initiatives, what can be attributed to payer-incentive programs is uncertain. Studies demonstrate that, of the 25 hospitals classified by the Centers for Medicare & Medicaid Services to be in the lowest decile of quality improvements, all still made significant progress in adhering to quality metrics after participation in the study. Financial rewards, however, were distributed based on a predetermined threshold established by the Centers for Medicare & Medicaid Services to be given only to participants who fell in the top 2 deciles. Penalties were incurred by the 51 hospitals that were within the bottom 2 deciles despite making substantial improvements. At such institutions, large minority communities and Medicaid populations comprise the patient populations. Other pay-for-performance schemes, such as employer-based purchasing, consumer health-spending accounts, and collaborative groups, were studied, with little data to support definite benefits.

Conclusions: Examining rates of improvement in adherence to pay-for-performance initiatives when determining how to distribute financial rewards should be studied alongside the current classification by absolute deciles. By rewarding rates of improvement, potential elimination of quality disparities for hospitals that serve large Medicaid and minority populations can be achieved, because such organizations are encouraged to invest in quality improvement as a result of substantial progress made. Although alternative strategies like employer-driven value-based purchasing and collaboratives seem promising, the long-term effects of such initiatives still need to be studied. Creating greater financial incentives for individual providers to participate in pay-for-performance programs for many years to come will remain a challenge. [*AHDB*. 2010;3(2):135-142.]

On March 1, 2001, the Institute of Medicine (IOM) released its report, "Crossing the Quality Chasm: A New Health System for the 21st Century," in response to alarming rates of medical errors that led to thousands of unnecessary deaths. The report called for changes within information technology (IT), payment policies, and the medical workforce. By stressing a "new paradigm for healthcare delivery," the IOM identified 15 medical conditions for which improvements could be made to improve the delivery of patient care.[1]

By examining publicly available medical records, a seminal study found that, on average, Americans received only 50% of recommended care for acute and chronic conditions.[2] Simple tasks like providing smoking-cessation counseling had only an 18.3% level of adherence. Dismal performance in diabetes management resulted in US providers achieving <25% the success of the UK's National Health Service (NHS) in reducing microvascular complications.[2] The Centers for Medicare & Medicaid Services (CMS) was called upon to quickly improve quality of care in the United States.

Dr Sura is a Resident in Diagnostic Radiology, St. Luke's-Roosevelt Hospital Center, New York, NY. Dr Shah is Assistant Professor of Medicine, NYU School of Medicine, and Associate Investigator, Geisinger Health System.

KEY POINTS

- In 2003, CMS established P4P initiatives to strengthen quality measures, improve patient outcomes, and maintain physician accountability.
- Such P4P programs offer incentives to hospitals, provider groups, and physicians based on adherence to specific composite metrics.
- This study raises the possibility that the quality composite scores promoted by CMS with P4P programs may also be in need of improvement.
- When treating minorities, providers are less likely to use the latest evidence-based guidance.
- For quality efforts to motivate doctors, it is becoming increasingly apparent that incentive programs have to represent a considerable fraction of their income.
- New P4P initiatives aim at creating market competition by establishing high and low performers. Others use collaborative practice groups to influence care.

This present study was undertaken to examine the quality initiatives endorsed by CMS in response to this effort. A literature search was conducted using the keywords "pay-for-performance," "quality improvement," "medical errors," and "physician incentive plans."

CMS Establishes P4P Initiatives

Responding to this call and aiming to strengthen quality measures, improve patient outcomes, and maintain physician accountability, CMS established pay-for-performance (P4P) initiatives at the end of 2003. The goals of the project were simple: if various healthcare stakeholders could not only demonstrate higher quality but also publicly report such results, financial incentives would follow. Through various initiatives and demonstrations, such as the Hospital Quality Incentive Demonstration (HQID) and Chronic Care Improvement Program, CMS developed quality metrics that targeted hospital facilities, physician groups, and individual providers to improve practice habits.

The purpose of HQID was to distribute financial rewards to hospitals that demonstrated high-quality performance on 34 quality measures established by CMS across 5 clinical conditions addressing acute care. After 3 years of using a nationally standardized set of quality measures to evaluate individual hospital performance, composite quality scores were calculated using an aggregate of adherence to all quality measures. A comparison was then made to a predefined threshold determined at the start of the project. For example, acute myocardial infarction (MI) was 1 of the 5 clinical conditions studied. Nine quality measures were intended to be met for any patient who was diagnosed with the condition. From the use of aspirin at arrival to prescribing beta-blockers at discharge, such metrics were closely followed.

Hospitals were then ranked according to their composite quality scores, with the first 10% placed in the top decile.[3] Hospitals in the top decile implementing the greatest number of quality measures would receive a 2% bonus at the end of the year. In the third year of the demonstration, hospitals that fell into the last 2 deciles would incur a 1% to 2% cut from their Medicare diagnostic-related grouping payments.[3]

In the medical community, such metrics can seem obvious when treating each condition individually. However, they can also quickly define and label a patient with the wrong diagnosis, leading to further tests and medications the patient does not need. For example, patients who have chest pain are often misdiagnosed with having an acute MI, when gastroesophageal reflux or musculoskeletal aches are the actual cause of their pain. Other times, the patient carries diagnoses that can confound the pathophysiology of the conditions, in which particular drugs or processes that CMS holds up to high standards are not indicated. For instance, beta-blockers, although listed as a quality metric to be given on arrival for all patients who present with acute MI, are not indicated in patients who also have a heart rate <60 beats per minute, systolic blood pressure <100 mm Hg, or moderate-to-severe stage C congestive heart failure (CHF).[4]

Aside from implementing quality measures, CMS established the Medicare Care Management Demonstration, which promoted the adoption of technology such as electronic medical records (EMRs). Bonus payments were awarded to practice groups whose members proceeded with IT implementation and used the EMR to collect data on meeting clinical quality measures.[5] Although many practices were convinced that the use of EMRs was applicable to the needs of their practices, the transition to that technology would have to occur slowly.[6] With a wide array of products, combined with a fast-paced, dynamic IT market, the uncertainty of how to integrate such software was a resonating theme.[6] Indirect and direct costs of implementation, accessibility, and daily use could make it difficult to justify an EMR application when a small practice struggles to simply meet its bottom line.

CMS Hospital Quality Incentive Demonstration

In November 2005, after 1 year of implementation of the HQID initiative, CMS conducted an internal review of the composite quality scores for 5 clinical conditions—acute MI, community-acquired pneumonia

(CAP), CHF, coronary artery bypass grafts, and hip and knee replacements.[3] The results demonstrated that improvement was made across all indicators by every participating hospital. From using prophylactic antibiotics before surgery to providing smoking-cessation assistance to patients with CAP, CMS participants boasted an average improvement of 6.6% in 22 of the 34 clinical indicators initially set forth. As had been promised, CMS paid more than $8.85 million in incentives to the highest 20% of performers. Hospitals in the ninth and tenth deciles faced reimbursement penalties for future care provided.[3]

In analyzing results through June 2007, the 255 hospitals that participated in the HQID project demonstrated a 17.3% increase over 5 years in median composite quality scores, a combination of clinical quality and outcome measures.[3] Median appropriate care scores—calculated by determining whether a patient received all possible care measures within a clinical area—also improved over 5 years by an average of 52.6% across all the 5 clinical conditions. When compared with nonparticipating hospitals, hospitals participating in the HQID project achieved 6.5% higher scores for 19 publicly reported quality indicators.[3]

Results of Independent Studies

After CMS completed its internal review, a number of studies tried to recreate these findings. At first glance, many seemed to confirm CMS's claims. In a study examining the additive benefit of P4P over public reporting of outcomes, hospitals participating in P4P programs not only made greater progress in 7 of the 10 individual performance metrics but also made large strides on all other composite measures of quality.[7] Although all baseline scores were higher after 1 year for those enrolled in P4P programs, the rate of improvement varied among those in the 5 quintiles studied.[7]

Hospitals in the lowest quintile at the start of the study made the largest gains in quality improvement. For instance, P4P-participating hospitals (N = 51) in the fifth and lowest quintiles had a 25.2% increase from an initial dismal 62.6% rate of adherence rate to quality metrics established in patients who presented with heart failure. For the same condition, the 51 hospitals in the first quintile actually dropped 0.1% from their previous adherence rate of 93.7%.[7] This trend was true across all 10 composite measures studied. Such "high-performing" hospitals experienced a "ceiling effect" on improvement, arguing that little improvement, if any, was left to be made at the start of the study.[7]

Achievements made by the lowest quintile or "low-performing" hospitals, however, were overshadowed by CMS's method of incentive distribution. CMS rewards absolute performance based on what decile a hospital lies in rather than on rates of improvement, so that the 102 participants who remained in the lowest 2 deciles were anticipated to receive a 1% to 2% financial penalty placed on their Medicare payments, despite making 29.2% of the 41.2% of total improvements achieved by all participants during the course of the 2-year study period. For hospitals that were penalized at the end of the study despite achieving success in reorganizing their process flows, many may not see the value of participating in P4P programs over the long-term.[7]

Another study discredited not just CMS's classification of low- and high-performing hospitals but the entire CMS initiative, citing a lack of control group for comparisons as a fatal flaw.[8] According to this analysis, independent hospitals outside the CMS initiative showed the same rate of improvement for all 9 of CMS's composite measures for patients admitted with acute MI.[8] Simply by adhering to their own quality initiative programs, the American College of Cardiology standards, or the American Heart Association guidelines, independent hospitals demonstrated equivalent gains.

Clinical outcomes (which have not been a focus of the CMS initiative) were also studied in great detail in addition to process measures. Mortality scores of CMS-participating hospitals were not incrementally greater than that of independent hospitals. P4P had little, if any, incremental impact on a patient's care or outcome during an acute MI, as measured by 14 process measures closely looked at in the study for patients with established non–ST-segment elevations on electrocardiogram.[8] Improvement in quality attributable to P4P was not as significant in the study (1.6% over 3 years[8] vs 4.3% unadjusted over 2 years,[7] as reported earlier by Lindenauer's analysis of CMS[7]), demonstrating the lack of concrete evidence linking P4P programs to better quality of care.[8]

Whether improvements made at CMS-participating hospitals can be attributed to P4P is debatable, but another question arises. Are the quality composite scores that CMS hopes to improve also in need of improvement? According to some researchers, analyzing the existing process flow and level of adherence to quality indicators for a particular hospital, practice, or department is an important first step before implementing performance metrics.

For example, emergency departments and urgent care clinics nationally have implemented many of CMS's guidelines for years. Performance metrics—such as treating an acute MI with aspirin or monitoring pulse oximetry in patients with CAP—already have high levels of adherence nationally (94.7% and 99.4%,

respectively).[9] Although the HQID-participating hospitals reported higher scores in adherence to 19 quality indicators compared with nonparticipating hospitals, nonparticipating hospitals may not see a strong need to achieve any incremental benefit if their levels of adherence are already high.

In contrast, Glickman and colleagues argue that focusing on the use of existing research networks, large quality-improvement registries, and further departmental analysis may be more beneficial.[9] Understanding the dynamic structures of each department (ie, the use of IT, incentive structures, culture, treatment algorithms, patient waiting times, and treatment durations) can potentially garner greater success in developing a framework centered on composite measures that follow.[9]

Although some composite measures achieve high levels of compliance nationally, others lack additional detail to aid the physician in improving the patient's clinical outcome. Merely documenting vital signs and oxygen saturation in the care of patients with CAP, as suggested by CMS, is not considered a higher standard of care if such data are not correlated with, for example, the patient's ventilator settings or arterial blood gas. Such correlations should be included when drafting the composite metrics that CMS aims to establish.[9]

Potential Disparities Created by P4P Programs

Disparities in medicine over the treatment of multiple conditions have been well documented.[10,11] When treating minorities, providers are less likely to use the latest evidence-based trends.[12-14]

Black Patients

Using Medicare data, researchers have examined whether hospital performance varies, after controlling for pertinent variables, as a function of the percentage of black patients the hospital serves.[15] In the care of patients with acute MI and CAP, a significant inverse relationship exists between hospitals' composite performance scores and the percentage of black persons these hospitals treat. Such hospitals are at increased financial risk for treating blacks under the CMS P4P program.

Although explanations seem vague as to why serving minority communities results in lower performance scores, some investigators have suggested that blacks are served by physicians who have no specialty board certifications and have reduced access to state-of-the-art technology.[15] Additional research must be conducted to determine why such a correlation is true. Is it a reflection of the physicians at such hospitals and their respective training, or is it the case complexity of the populations they treat?

Regardless of the reasons, with lower performance scores, hospitals that serve larger black populations will find it even more difficult to invest in CMS-suggested process flows that strive to improve performance if CMS classifies them in the lower deciles and penalizes them financially. With poor performance results that are publicly reported, such hospitals will likely blame the minority patients they serve, citing a strong correlation with their performance and case mix, giving rise to administrators seeking healthier, higher socioeconomic patients to serve.

Medicaid Populations

Current P4P schemes are not just affecting black populations but all patients covered by Medicaid. Safety-net hospitals that serve a large percentage of Medicaid patients are finding it increasingly difficult to achieve performance metrics enforced by CMS.

Of the 4464 hospitals that participated in a recent study, those that treat a higher percentage of Medicaid patients had comparatively worse performance scores and made significantly smaller improvements over time, classifying them as low performers.[16] Hospitals caring for fewer Medicaid patients had an average performance gain of 39% and were subsequently classified as top performers. Top-performing hospitals serving a high percentage of Medicaid patients who presented with an acute MI had nearly a 75% decrease in performance from 2004 to 2006. At the same time, top-ranked hospitals serving a low percentage of Medicaid patients for the same composite measure, acute MI, had increased gains between 13.6% and 19.7% in performance. These patterns held true across 3 other P4P conditions for which CMS seeks to make quality improvements.[16]

With such results, hospitals serving primarily Medicaid patients would fall significantly below the threshold, incurring financial penalties rather than receiving the bonuses necessary to allow these facilities to make infrastructural changes.

Potential Correction

It is certain that care for underserved patients concentrates at low-performing hospitals. With revenue margins for such hospitals already in the red, it becomes difficult to shift costs onto payers, while continuing to deliver healthcare. By restructuring CMS' performance schemes to look at rates of improvement along with absolute numbers, CMS can potentially correct some of the disparities, allowing P4P financial gains to be accrued at such hospitals over time. Also, by adjusting for patient case mix and treatment opportunity mix, CMS might observe moderate changes that affect hospital performance, rankings, and eligibility for P4P financial benefits for those hospitals currently ranked lowest.[17]

P4P and Its Use Outside of CMS

In an effort to instill quality initiatives among healthcare providers, CMS does not remain alone. Major corporations in multiple industries have joined efforts to improve quality of care.

General Electric, Partners Healthcare, Tufts Health Plan, the Lahey Clinic, and other employers in Massachusetts have developed Bridges to Excellence, a program initiative that rewards physician offices up to $55 per patient annually for quality improvements made in IT, care-coordination systems, and decision support teams. In addition, physicians who are recognized as an American Diabetes Association provider for implementing exceptional diabetic care are paid rewards of up to $100 per patient.

Six California health plans under the collaboration called Integrated Healthcare Association seek to improve standards of care with the use of consolidated physician performance scorecards.[18]

The Leapfrog Group has also made significant strides in mobilizing employer purchasing power, working with corporations to encourage transparency, as well as to financially reward hospitals that have a demonstrated record of high-quality care.[19]

Anthem Blue Cross and Blue Shield in New Hampshire launched an incentive program to improve quality in 1999, focusing on rewarding practices for implementing preventive measures, such as cancer screening, well-child examinations, and childhood immunizations. Under this program, physicians who ranked in the top 25% of adherence to metrics relative to other providers within the network received an additional $20 per patient per year. In 2002, physician practices received up to $12,062 in bonuses for practicing higher standards of care in the previous year.[18]

Yet for quality efforts to motivate doctors, it is becoming increasingly apparent that incentive programs have to represent a considerable fraction of their income. Too many P4P initiatives prevent providers from engaging in such programs; the costs to enact the metrics outweigh the financial merits received from them. For example, under the current Bridges to Excellence program, incentives only cover 1% of the patient's expenses in a general internist's office.[18] This equates to just more than $1200 of additional revenue for the entire year in a practice with 2300 patients.[18] With continued Medicare cuts (5.4% in 2002), physicians may be forced to adopt other practice alternatives, such as increasing patient volume, rather than implementing payment incentives.[18]

New Strategies to Improve P4P

Currently, P4P initiatives are aimed at creating market competition by establishing high and low performers. The theory behind the model is to motivate groups to achieve above and beyond the threshold of their fellow competitors. Others, however, have formulated different methods to encourage quality improvement. Coined "collaborative" practice groups, they try to influence care broadly by bringing community health centers together to learn and disseminate quality-improvement techniques.[20] The following chronology is used:

- Come to learn
- Implement
- Intervene
- Share results.

E-mail updates, monthly reports, and continuous feedback are also integral components to such a collaborative effort.

Impressive results have been reported.[20] Intervention centers that participated in one such collaborative made significant gains in composite indicators. From prevention to disease management, the participating centers boasted up to a 6.2% higher rate of improvement in adherence to such indicators compared with centers that did not participate.[20] One significant limitation to the study's results, however, was that no significant improvement in outcomes was seen. Positive long-term results from the collaborative are yet to be determined.

Another innovative strategy has been to involve patients directly by allowing consumers of healthcare to examine quality data and make decisions for themselves. Such a system has been proposed, aligning performance initiatives to that of consumer purchasing power. First, a healthcare spending account is established. The consumer has access to data that examine efficiency and quality among providers and hospitals. Assuming that consumers then make economical choices and are motivated to seek physicians who provide high-quality, low-cost, and efficient care, they are later rewarded with tax breaks for unspent funds that can be used for future unexpected healthcare expenditures.

Problems arise with fixed consumer accounts. Consumers often like having their healthcare decisions made by others, such as their employers. For employers, wide coverage variability, with different market prices, makes assigning a fixed dollar amount to spending accounts difficult. Creating access to such quality ratings becomes problematic. How much information does one make transparent? Providers have been staunchly opposed, mainly as a result of the high direct costs they face in collecting data, as well as quality measures that are undefined and variable from practice to practice.

Yet some initial studies show that linking financial incentives to individual empowerment based on dis-

seminated quality of care data has its benefits. In one program, patients and families who selected organ transplant centers on the basis of quality and costs, and were later rewarded financially for doing so, had a high level of measured satisfaction once their surgery was complete.[21] Minnesota's Buyers Health Care Action Group released data to consumers in the market regarding health plans that were low in cost yet high in quality ratings based on satisfaction surveys completed by enrollees. Enrollment increases ranged from 15% to 57% among the top 4 plans once the information was made public.[22]

In a Merck/Medco-sponsored study conducted in 2001, a more impressing approach was taken, penalizing consumers with a 10% increase in premiums if they did not select lower-cost, yet equally efficacious, medications within similar drug groups.[23] Financially incentivizing the consumer worked; by 2005, the company boasted a substantial increase in the use of more cost-effective drugs. Each 1% increase in the use of generics yielded an estimated $200 million in savings for the Merck/Medco plan sponsors and members.[23]

While incentivizing consumers may prove beneficial, pharmaceutical companies still struggle to convert efficacy in clinical trials to effectiveness in clinical practice. For example, although new medications have improved survival in patients who develop heart failure, the improvement in physician use is not as dramatic as drug companies would like to see. Despite clinical trials all concluding that angiotensin-converting enzyme inhibitors, beta-blockers, and aldosterone antagonists have some beneficial role in a patient who has suffered an acute MI, underutilization of these therapies in clinical practice reflects no significant change in epidemiologic data.[24]

Employers' Use of Value-Based Purchasing Strategies

In 2007, 158 million nonelderly Americans were covered under employer-based health insurance, with employers becoming the leading source of health insurance in the United States. With annual premiums averaging 5% higher than the previous year and 120% higher since 1999 after accounting for inflation, one would think firms would look to sponsor P4P initiatives and value-based plans with the goal of keeping increasing healthcare costs from translating into higher premiums.[25] The Leapfrog Group and Bridges to Excellence have led many of the efforts to encourage companies to become active participants in comparing quality when selecting plans. Through the promotion of hospital safety measures profiling as well as providing tools to help employees to make healthcare decisions, such groups have assisted employers in providing greater transparency for their employees' healthcare decision-making.[26]

Other priorities, however, are of greater concern to executives. Across 41 US markets, firms rank geographic distribution and premium rates as the most important characteristics when choosing health plans 85% of the time.[26] Metrics ostensibly reflective of quality of care are given far less weight. Whereas 65% of employers are vigilant of patient satisfaction indicators for health plans, few reported making such information transparent to their employees.[26] The use of report cards, bonuses in contractual plans, and premium contribution to encourage employees to choose better plans, are provided to employees making healthcare decisions by firms only 23% of the time.[26]

Although larger employers are engaging the concept of value-based purchasing more than their smaller counterparts, the concepts of examining IT infrastructure (shown in multiple studies to reduce errors and improve quality of care[27,28]) with physician and medical group quality data were studied less than 20% of the time among all employers.

Because of a lack of concrete cost-benefit analysis, 65% of employers rarely engage in value-based purchasing primarily. Many small- or medium-sized businesses incorrectly translated higher value to mean higher premiums. Others could not justify spending the additional premiums now on attractive health plans when research regarding the intangible effects on workforce productivity, benefit to cost-savings, and the ability to attract and retain employees is inconclusive.[26]

Conclusions

In 2003, England's NHS outlined a proposal instituting 76 quality initiatives to aid 10 clinical domains of care. The plan accompanied a substantial increase in payments to providers who adhered to such targets.[29] Did CMS and private payers simply try to apply NHS' quality initiatives to their broken system too hastily? By developing their own P4P initiatives, CMS also sought to address increasing costs and poor performance by linking reimbursement to adherence to standards purported to measure the quality of the healthcare being provided by physicians and hospitals.

Initial studies demonstrated improvement in quality for hospitals participating in the P4P initiative. However, once comparisons were made to hospitals not participating in the CMS project, results were variable regarding differences in quality composite scores between the 2 groups. Research has shown that small financial gains are simply insufficient as currently provided and are not enough to motivate providers and hospitals to do better. Problems arose with how to distribute payments, as well as whether improvements of some of

the metrics set forth were needed.

Hospitals categorized as low performing as a result of their current standing in a low decile made improvements but not enough to be classified as top performers, making it difficult to continue to serve their patient and payer mix, especially if the majority of their patients are covered only by Medicaid. By examining rates of improvement, along with absolute categorization in a top decile, distribution of financial incentives can motivate all participating and potentially eliminate disparities.

Other alternatives, such as collaborative, consumer-directed spending accounts, pharmaceutical initiatives, and employer-driven initiatives, have shown some success, but more data are needed. Although P4P programs in some form likely are here to stay, keeping providers invested in closing gaps in disparity, while instituting plans that work, will be challenging.

Disclosure Statement

Dr Shah receives unrestricted research grants from AstraZeneca, Merck, Pfizer, and Roche.

References

1. Washburn ER. Fast forward: a blueprint for the future from the Institute of Medicine. *Physician Exec*. 2001;27:8-14.
2. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. *N Engl J Med*. 2003;348:2635-2645.
3. Premier. Centers for Medicare & Medicaid Services (CMS)/Premier Hospital Quality Incentive Demonstration Project. April 13, 2006. http://premierinc.com/quality-safety/tools-services/p4p/hqi/hqi-whitepaper041306.pdf. Accessed January 14, 2009.
4. Klapholz M. Beta-blocker use for the stages of heart failure. *Mayo Clinic Proc*. 2009;84:718-729.
5. Centers for Medicaid & Medicare Services. Medicare "Pay for Performance (P4P)" Initiatives. Press release. January 31, 2005. www.cms.hhs.gov/apps/media/press/release.asp?Counter=1343. Accessed January 13, 2009.
6. Hampton T. Groups push physicians and patients to embrace electronic health records. *JAMA*. 2008;299:507-509.
7. Lindenauer PK, Remus D, Roman S, et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med*. 2007;356:486-496.
8. Glickman SW, Ou F, DeLong ER, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA*. 2007;297:2373-2380.
9. Glickman SW, Schulman KA, Peterson ED, et al. Evidence-based perspectives on pay for performance and quality of patient care and outcomes in emergency medicine. *Ann Emerg Med*. 2008;51:622-631.
10. Ferguson JA, Tierney WM, Westmoreland GR, et al. Examination of racial differences in management of cardiovascular disease. *J Am Coll Cardiol*. 1997;30:1701-1713.
11. Mortensen EM, Cornell J, Whittle J. Racial variations in processes of care for patients with community acquired pneumonia. *BMC Health Serv Res*. 2004;4:20-26.
12. Sonel AF, Good CB, Mulgund J, et al. Racial variations in treatment and outcomes of black and white patients with high-risk non-ST-elevation acute coronary syndromes: insights from CRUSADE (Can Rapid Risk Stratification of Unstable Angina Patients Suppress Adverse Outcomes with Early Implementation of the ACC/AHA Guidelines?). *Circulation*. 2005;111:1225-1232.
13. Peterson ED, Wright SM, Daley J, Thibault GE. Racial variation in cardiac procedure use and survival following acute myocardial infarction in the Department of Veterans Affairs. *JAMA*. 1994;271:1175-1180.
14. Kaul P, Lytle BL, Spertus JA, et al. Influence of racial disparities in procedure use on functional status outcomes among patients with coronary artery disease. *Circulation*. 2005;111:1284-1290.
15. Karve AM, Ou FS, Lytle BL, Peterson ED. Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *Am Heart J*. 2008;155:571-576.
16. Werner RM, Goldman LE, Dudley RA. Comparison of change in quality of care between safety-net and non-safety-net hospitals. *JAMA*. 2008;299:2180-2187.
17. Mehta RH, Liang L, Karve AM, et al. Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. *JAMA*. 2008;300:1897-1903.
18. Epstein AM, Lee TH, Hamel MB. Paying physicians for high-quality care. *N Engl J Med*. 2004;350:406-410.
19. The Leapfrog Group. About us. www.leapfroggroup.org/about_us. Accessed January 26, 2009.
20. Landon BE, Hicks LS, O'Malley AJ, et al. Improving the management of chronic disease at community health centers. *N Engl J Med*. 2007;356:921-934.
21. Galvin R, Milstein A. Large employers' new strategies in health care. *N Engl J Med*. 2002;347:939-942.
22. Christianson J, Feldman R, Weiner JP, Drury P. Early experience with a new model of employer group purchasing in Minnesota. *Health Aff (Millwood)*. 1999;18:100-114.
23. AllBusiness. Merck-Medco Achieves 80 Percent Generic Substitution Rate for Prozac in First Week; Plan. August 21, 2001. www.allbusiness.com/health-care/health-care-regulation-policy-health/6112721-1.html. Accessed November 1, 2009.
24. Sackner-Bernstein J. Reducing the risks of sudden death and heart failure post myocardial infarction: utility of optimized pharmacotherapy. *Clin Cardiol*. 2005;28 (suppl 1):19-27.
25. The Henry J. Kaiser Family Foundation. Kaiser Commission on Medicaid and the Uninsured. Health Insurance Coverage in America: 2006 Data Update. October 2007. www.kff.org/uninsured/upload/2006_DATA%20_UPDATE.pdf. Accessed March 3, 2010.
26. Rosenthal MB, Landon BE, Normand SL, et al. Employers' use of value-based purchasing strategies. *JAMA*. 2007;298:2281-2288.
27. Blumenthal D, Glaser JP. Information technology comes to medicine. *N Engl J Med*. 2007;356:2527-2534.
28. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med*. 2006;144:742-752.
29. Shekelle P. New contract for general practitioners. *BMJ*. 2003;326:457-458.

STAKEHOLDER PERSPECTIVE

The Many Challenges of Pay-for-Performance Programs

PAYERS: Drs Sura and Shah focus their article on the benefits of pay-for-performance (P4P) programs in the hospital setting, while also noting some problems related to office-based healthcare. Although hospitalization represents a large cost for relatively brief episodes of care, the majority of patients receive their care in the physician's office setting. Office-based P4P programs have some of the same potential benefits as in the hospital, but they have been associated with even greater concerns.

As healthcare costs have increased faster than other costs, healthcare payers and purchasers continue to search for mechanisms to ensure high-quality, cost-efficient care. Such value-based purchasing (VBP) has allowed these third parties to transition from passive payers to active purchasers of care. Those involved in healthcare reform have identified 4 cornerstones upon which quality improvement activities can be based. One of these involves the adoption of incentives to promote high-quality, cost-effective care. The Deficit

Continued

STAKEHOLDER PERSPECTIVE *(Continued)*

Reduction Act of 2005 focused on VBP and P4P efforts. Executive Order 13410, issued in August 2006 by President Bush, recommended P4P models of reimbursement. In 2007, then–Health and Human Services (HHS) Secretary Michael Leavitt encouraged major healthcare purchasers to adopt VBP. The Medicare Improvements for Patients and Providers Act of 2008 required HHS to develop a plan to transition physicians to VBP/P4P. Private payers have rapidly implemented many P4P programs.

Some initiatives, such as CMS's Physician Quality Reporting Initiative (PQRI), reward physicians' voluntary reporting with monetary compensation. Other initiatives levy fines for hospital nonreporting. Private-sector programs frequently reward physicians for meeting not only quality-of-care goals but also nonclinical patient satisfaction. Private payer P4P programs often purport that their goal is to improve quality, but the process can be manipulated such that the primary goal is to reduce cost. With more than 100 P4P programs in the marketplace, providers have a very difficult time meeting the reporting requirements, which vary significantly based on the clinical measures, quality metrics, and the patient population targeted. Although it may be difficult to measure quality outcomes of care, many plans have substituted easier process measures instead of outcome measures; depending on the data, it may be very difficult to attribute care or lack thereof to a specific provider.

The goals of P4P programs are similar but their incentives vary significantly. Some P4P programs award modest bonuses to providers who meet program objectives; others offer bonuses only to providers who document improvement in meeting metrics; yet others reward only top performers. Other programs offer consumers reduced copayments, coinsurance, deductibles, or premiums for using only providers identified as "high quality." Newer programs reward care coordination models, such as implementing a medical home or accountable-care organizations.

Among the greatest challenges facing purchasers and public/private payers in implementing P4P/VBP programs is ensuring the validity and credibility of the metrics. Many metrics are based on treatment guidelines, and many of these guidelines, as Dr Jerome Groopman has elucidated,[1] are conflicting and are often produced by entities with significant potential for conflict of interest. Risk adjustment is also essential, because patients and physicians are not truly fungible entities. CMS's PQRI and other state reporting experiments have experienced significant problems.

Because of flaws in these programs, most have not produced the desired results. Despite extensive and frequent physician feedback about them, many programs continue to move toward public disclosure of inherently flawed data. Simple disclosure of these flaws, however, is insufficient—correction is essential. Programs in Massachusetts, New York, Texas, and Washington have publicly released flawed data, resulting in litigation and significant monetary settlements against payers.

PROVIDERS/PATIENTS: The goal of such quality feedback programs should be to educate physicians to assist them in practice management and provision of quality care. Paramount to the success of such programs is reliable, verifiable data; however, almost every state/federal program has experienced serious data accuracy problems. Inaccurate data can increase the risk of unintended consequences, mislead patients, harm a physician's reputation, and increase physician distrust in P4P participation. Without clinical data, claims data alone are clearly insufficient to document the full spectrum of specific patient care.

Patient adherence to recommended intervention must also be considered. Physicians must be given the opportunity to review the accuracy of the patient attribution and the clinical data. Harmonization of the metrics being studied must occur across all payers. Physicians must also be given the right to appeal and correct inaccuracies before any data or conclusions are released to the public. Publishing poor quality measures, and making access to physicians, dependant on potentially inappropriate and/or poorly attributed metrics, make the information relatively meaningless to those aware of the lack of validity, and potentially damaging to physician practices, if patients seek care based on this poorly collected and attributed data. Because of these many difficulties, physician participation in such flawed processes should remain voluntary until these problems are fully rectified.

1. Groopman J. *How Doctors Think*. New York, NY: Houghton Mifflin; 2007.

Steven T. Kmucha, MD, JD
Medical Director, Balance and Mobility Center at Seton Medical Center, CA
Consultant, California Medical Association–Committee on Quality Care, and The Doctors Company
Adjunct Clinical Instructor, Stanford University School of Medicine